PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Sedation versus general anaesthesia in endovascular therapy for anterior circulation acute ischemic stroke: the multicentre randomised controlled AMETIS trial study protocol
AUTHORS	Chabanne, Russell; Fernandez-Canal, Charlotte; Degos, Vincent; Lukaszewicz, Anne-Claire; Velly, Lionel; Mrozek, Segolene; Perrigault, Pierre-François; Molliex, Serge; Tavernier, Benoit; Dahyot-Fizelier, Claire; Verdonk, Franck; Caumon, Elodie; Masgrau, Aurélie; Begard, Marc; Chabert, Emmanuel; Ferrier, Anna; Jaber, S.; Bazin, Jean-Etienne; Pereira, Bruno; Futier, Emmanuel

VERSION 1 – REVIEW

REVIEWER	Judith Dinsmore St Georges Hospital, London, UK
REVIEW RETURNED	09-Nov-2018

GENERAL COMMENTS	The authors present the AMETIS Trial study protocol: a prospective multicentre, randomised clinical trial including 270 patients and comparing general anaesthesia and sedation during intra-arterial treatment for anterior circulation stroke. The Primary outcome is a composite of functional independence at 3 months and absence of medical complication occurring by day 7. The study began in August 2017. It is an important and clinically relevant topic. The efficacy of endovascular treatment in patients with anterior circulation stroke due to large artery occlusion is firmly established. Previous studies have suggested that endovascular treatment under general anaesthesia (GA) is associated with a worse outcome. However, as patients with more severe stroke and comorbidities are more likely to receive GA there is potential for confounding by indication. As such there is a need for large scale prospective randomised controlled trials. The protocol is generally well written however I have a few comments: Introduction: This lays out the background. However, what most agree is not just that a multicentre RCT is needed but that this is large scale. They are aiming to recruit only 270 patients. In addition, The CANVAS trial is currently ongoing in China. This is a prospective randomised equivalence trial investigating the effects of GA versus CS on outcome using mRS at 90 days. This is aiming to recruit 635 patients.

	Methods: In terms of inclusion/ exclusion criteria, there is no mention of allergies/ previous problems with GA. What about medical co-morbidities? There is no mention of NIHSS score in terms of inclusions / exclusions. The authors do state later in the discussion that no minimum or maximum NIHSS is recommended in order to achieve a pragmatic investigator based approach however most previous trials have included NIHSS as inclusion criteria as does the ASA /AHA guidelines that they quote for use. Premorbid mRS > 1 as an exclusion rules out many who might benefit. ASA /AHA say that it may be reasonable to proceed in these patients. This does not fit with a pragmatic trial without exclusions based on NIHSS. In terms of interventions, there is no protocol for either the general anaesthesia group or the CS group. I can understand why they are doing this, but it might make interpretation of results difficult as there may be a wide variation in practice. Will depth of anaesthesia monitoring be used – at least if a similar depth of anaesthesia was achieved there would be some standard to compare? Conversion to GA is recommended for coma, loss of protective airway reflexes, respiratory failure. Both the later would be difficult to assess intra-procedure and I presume they mean as a consequence of over-sedation. Otherwise they should be exclusion criteria. There is no mention of patient agitation or inability to co-operate? Although haemodynamic control and carbon dioxide targets have been set, there is no mention of blood glucose or other important physiological targets. The primary outcome is a composite of functional independence at 3 months (mRS 0-2) and absence of medical complication by day 7. There is little explanation as to why this composite score is being used and how this will be performed. I am not sure of the benefit of using this. Most of the previous work has used mRS at 90 days. I am not aware of any previous studies using this particular approach. Who will perform the 90-day mRS assessment? Why medical complications by day 7? It would be helpful if the authors could explain their reasoning. For the secondary outcomes they include successful recanalization but there is no mention of infarct volume – an important variable in terms of outcome. There is no definition of what constitutes hypotension – how low and how long? There is not enough information on the specific important time points which will affect outcome. Door to groin puncture time – is this hospital door or radiology suite? Stroke onset time? Time of induction of anaesthesia / sedation, duration of procedure? Why day 7 for medical complications and unexpected ICU admission For sample size estimation, the authors have used 5 previous studies on anterior circulation stroke. However, all of these used a primary outcome of mRS and so they have extrapolated composite scores. I am not sure how reliable this will be.

REVIEWER	Adeel Ilyas, MD University of Alabama at Birmingham, Birmingham, AL, USA
REVIEW RETURNED	23-Nov-2018

GENERAL COMMENTS	The authors have outlined a protocol designed to determine outcome of endovascular mechanical thrombectomy for anterior circulation acute ischemic stroke under two different anesthesia modalities. The composite primary outcome included medical complications, though I believe these are both medical and surgical in nature. Among these complications, pneumonia is listed. Patient may have pre-morbid conditions that predispose them to hospital acquired pneumonia (e.g. chronic obstructive pulmonary disease), or they may present to the hospital with pneumonia. These need to be accounted, and are not listed within the data collection parameters (Supplementary File 1: At randomization). Furthermore, the definition of cardiogenic acute pulmonary edema and how it is distinguished from pulmonary edema secondary to fluid overload is not stated clearly as patients with poor ejection fracture are likely to develop pulmonary edema secondary to fluid overload. This outcome is too nuanced and should be removed or clearly defined. In the Abstract, Introduction: EMT is the standard of care for anterior circulation AIS secondary to emergent large vessel occlusion in patients who qualify (i.e. not all types of anterior circulation AIS). Overall, the syntax and word choice can be improved, for example: 1. In the Abstract: Introduction, the second sentence ("To ensure patient ... proposed.") is awkward. 2. In the Introduction: Background and rationale, the sentence ("Notably, anesthetic management ... general anesthesia (GA).") is poorly written and does not effectively convey equipoise between to anesthetic modalities.

REVIEWER	Lene Theil Skovgaard Dept. of Public Health, Section of Biostatistics, University of Copenhagen, Denmark
REVIEW RETURNED	15-Jan-2019

GENERAL COMMENTS	The design is adequate and well described. The two primary outcomes are described several times but I am still not quite sure how they are defined and how they will be treated. It seems that "Functional independence" is an ordinal variable which can take the values 0,1 and 2, and the analysis of this is presumably described on page 17, line 31 ff. A Chi2 test or Fishers exact test seems appropriate, but where does the Poisson distribution come from? There is no count variable here...? Or does this refer to the second primary outcome, called "absence of medical complications..."? Perhaps the medical complications are being counted? Although from page 6, it seems that this variable is thought to be dichotomous. On page 34, the Poisson analysis is mentioned again, this time as an analysis of "rate data", what is that? Even so, results will be expressed as relative risks, a term that relate to logistic regression. I am confused. The sample size calculations are OK, although a 10% loss to follow-up will leave only 270*0.9=243 patients, i.e. not quite 124 in each group. Also, the sample size may not be sufficient for the adjusted analyses.

	The covariates are described adquately. In general, I prefer the term "multiple regression" to "multivariable analysis" or just mentioning that the models contain multiple covariates.

VERSION 1 – AUTHOR RESPONSE

REVIEWER # 1: We thank Doctor Dinsmore for her pertinent reviewing and will try to answer her interrogations:

• Introduction: This lays out the background. However, what most agree is not just that a multicentre RCT is needed but that this is large scale. They are aiming to recruit only 270 patients. In addition, The CANVAS trial is currently ongoing in China. This is a prospective randomised equivalence trial investigating the effects of GA versus CS on outcome using mRS at 90 days. This is aiming to recruit 635 patients.

Concerning the sample size, we agree that, despite multicenter, the AMETIS trial will not be a very large study. As you mentioned, the ongoing CANVAS non-inferiority trial is planning to enroll 635 patients in China(1). Indeed, sample size estimation on this recent research topic appears difficult related to a still quite exploratory question with few high-quality data. For example, we could also cite others ongoing trials with the modified Rankin scale (mRS) as a primary endpoint like the GASS study (NCT02822144), a multicenter superiority trial of conscious sedation (CS) versus general anaesthesia (GA) planning to include 350 patients or the SEGA study (NCT03263117), a monocentric superiority trial of GA versus CS planning to include 260 patients. As we mentioned in the discussion section for possible limitations of our protocol and the choice of a composite outcome instead of the mRS as the primary endpoint: “The effect size that we could expect on functional independence at 3 months is probably far less than thrombectomy on its own. Based on actual literature, SIESTA trial found dramatically decreased functional independence associated with CS with only 18% of mRS 0-2 compared to 37% in GA.(2) 18% of patients being independent is far less than in thrombectomy trials where it barely represents controlled groups (intravenous thrombolysis alone).(3-8) With these proportions, 240 patients would have been necessary to demonstrate a statistical difference with a beta power of 90% but we could expect important centre effect in SIESTA trial. On the contrary, ANSTROKE trial didn’t find any difference between groups, with functional independence in respectively 42 and 40% of patients between GA and CS.(9) Based on these 2 trials, functional independence could be obtained in roughly 40% of patients under GA. Providing a 20% variation in positive or negative effect on functional independence, more than 1000 patients would be required with a 80% beta power. An anaesthesia size effect of more than 20% appeared unrealistic. » Our study is a multicentre trial but is not intended to definitely conclude about the question and it will be very interesting to put together the results of the aforementioned studies with ours. We made precisions on the discussion part of our manuscript: “Concurrent ongoing trials with day 90 mRS as a primary outcome are planning to recruit 635 patients to demonstrate non-inferiority between CS and GA,(1)(350 patients to demonstrate superiority of CS vs GS (NCT02822144) or 260 patients to demonstrate superiority of GA vs CS (NCT03263117).”

• Methods: In terms of inclusion/ exclusion criteria, there is no mention of allergies/ previous problems with GA.

Concerning the inclusion/exclusion criteria we agree that we do not list all specific factor affecting anaesthesia care especially allergy and anteriority with anaesthesia. Nevertheless, these elements are the core principles rules of every professional anaesthesia provider that is mandatory in France. Every investigator that has the ability to include a patient in the study is a certified senior Anaesthetist in charge of the patient. Furthermore, related to the emergency fashion of the intervention with possible aphasic patients without proxies, it could be difficult to have a complete medical history.

Every SAE and SUSAR will be declared and an interim security analysis is planned. We reinforce this aspect in the manuscript:
“Interventions : Patients eligible for inclusion will be randomly assigned to CS or GA after a routine medical anaesthetic emergency evaluation has been made by a certified senior Anaesthesiologist. As required by French law, all contraindications and/or known allergy to anaesthetics will be registered.”

- Methods: What about medical co-morbidities?

Concerning comorbidities, as mentioned in the supplementary file 1: AMETIS trial data collection, medical history of hypertension, renal failure, cardiac failure, diabetes mellitus, alcohol abuse and active smoking will be collected at inclusion. Moreover, stroke complicating another acute illness or postoperative stroke are exclusion criteria in order to avoid confounding factors in the primary outcome measure.

- Methods: There is no mention of NIHSS score in terms of inclusions / exclusions. The authors do state later in the discussion that no minimum or maximum NIHSS is recommended in order to achieve a pragmatic investigator based approach however most previous trials have included NIHSS as inclusion criteria as does the ASA /AHA guidelines that they quote for use.

As Doctor Dinsmore mentioned, we made the choice not to integrate the NIHSS score as an inclusion/exclusion criteria. In our pragmatic design, indication for thrombectomy will be made by the attending senior vascular neurologist and neuroradiologist based on actual recommendations that is at least a NIHSS score of 6 whatever the affected side is. In certain situations, as mentioned in the recommendations, thrombectomy could even be indicated when NIHSS is < or = 6.(10) Thereby, when the procedure is indicated, whatever the NIHSS is, the question of the optimal anaesthetic strategy still exists and deserves evaluation. Notwithstanding, a stratification on the NIHSS score is planned in order to a have homogeneous NIHSS repartition between groups. We added in the discussion: “Notably, despite published trials mentioned NIHSS limits as inclusion/exclusion criteria, providing thrombectomy is indicated based on actual recommendations, the optimal anaesthetic strategy deserves evaluation whatever the NIHSS is.”

- Methods: Premorbid mRS > 1 as an exclusion rules out many who might benefit. ASA /AHA say that it may be reasonable to proceed in these patients. This does not fit with a pragmatic trial without exclusions based on NIHSS.

We agree that thrombectomy in premorbid mRS>1 might benefit to patients, which is corroborated by ASA/AHA guidelines.(10) Nevertheless, considering the difficulty to evaluate premorbid mRS in emergency condition, we feared to include dependent patients which could strongly affect the primary outcome especially day 90 functional independence defined as a 0-2 mRS. This strategy was adopted by others in study about thrombectomy. We integrated this aspect in the limitations of the study:
“Third, although thrombectomy might benefit to patients with premorbid mRS>1, we excluded these patients since evaluation may be somewhat difficult in emergency conditions. This strategy was adopted by others.(5-7,11)”

- In terms of interventions, there is no protocol for either the general anaesthesia group or the CS group. I can understand why they are doing this, but it might make interpretation of results difficult as there may be a wide variation in practice. Will depth of anaesthesia monitoring be used – at least if a similar depth of anaesthesia was achieved there would be some standard to compare?

We agree that the protocol for either the GA group or the CS group is not standardized. As mentioned in the limitation section of the manuscript, we choose not to protocolize GA and CS since “no data demonstrate that a drug is better than another even if modulation of CBF could be variable. However, the protocol requires strict objectives for systolic blood pressure and “normal” blood carbon dioxide tension in GA group.(12,13) Drugs and dose will be monitored. » Moreover, a rigid and standardized protocol might also be open to criticism.14 Further studies are required to explore whether specific anaesthetic protocols are associated with difference in outcomes. Some already published and ongoing trials imposed specific anaesthetic protocols and comparison will be interesting.(1,9,15) We agree that monitoring the depth of anaesthesia would facilitate the interpretation of study findings. However, there is currently no recommendation for the monitoring of the depth of anaesthesia in France. Furthermore, disappointing results were recently published.(16) Despite its use is possible,

frontotemporal electrodes and cable could interfere with the procedure concerning quality of radiographic images in this specific setting. CS depth will be evaluated by a dedicated clinical scale.

- Conversion to GA is recommended for coma, loss of protective airway reflexes, respiratory failure. Both the later would be difficult to assess intra-procedure and I presume they mean as a consequence of over-sedation. Otherwise they should be exclusion criteria. There is no mention of patient agitation or inability to co-operate?

As mentioned in the protocol, CS conversion to GA is recommended for severe agitation, coma, loss of airway protective reflexes, respiratory failure and incoercible vomiting. We believe that all these complications could be assessed intra-procedure. Coma, loss of airway protective reflexes and respiratory failure could be due to over-sedation or secondary per-procedural neurological deterioration. Agitation or patient inability to cooperate are not exclusion criteria, providing it is not associated with altered vigilance, since these 2 conditions could be managed by both CS or GA. Aphasia, which is frequent in stroke, compromises patient cooperation. Indeed, these could be factors of CS failure but it has to be demonstrated and it will be evaluated.

- Although haemodynamic control and carbon dioxide targets have been set, there is no mention of blood glucose or other important physiological targets.

Patients will be managed by the vascular neurology team prior, during and after the procedure with a senior neurologist and a specialized nurse. Blood glucose is systematically assessed before thrombectomy with an extensive blood analysis and treated as needed with monitoring according to institutional stroke management protocols. We didn't recommend any specific glycemic control related to the study.

- The primary outcome is a composite of functional independence at 3 months (mRS 0-2) and absence of medical complication by day 7. There is little explanation as to why this composite score is being used and how this will be performed. I am not sure of the benefit of using this. Most of the previous work has used mRS at 90 days. I am not aware of any previous studies using this particular approach. Who will perform the 90-day mRS assessment? Why medical complications by day 7? It would be helpful if the authors could explain their reasoning.

We will use a composite outcome as the primary endpoint. We would like to apologize for the lack of clarity explaining this choice and how it will be measured. Components of the composite, except mRS, will be assessed at day 7 (or at hospital discharge if prior) by trial or trained research staff blinded to the allocation group using patient medical records. Day 90 mRS will be centrally evaluated by phone by a certified blinded assessor. Structured mRS interview will be provided to the patient whenever possible. Otherwise, proxy or healthcare provider of the patient will be asked. We choose a composite outcome in order to 1) combine related relevant perioperative outcomes representing different aspects of a unique pathophysiological process modifiable by the intervention 2) increase event rate to improve statistical power and 3) evaluate every dimension of perioperative care possibly influenced by the intervention that could have impact at a patient or society level (e.g. prolongation of hospitalization). Although important, mRS imperfectly reflects the influence of perioperative anaesthesia care on outcomes. For example, GA could prevent patient movement and avoid vessel perforation/dissection but pneumonia or myocardial infarction could be increased. These events impact significantly perioperative care without necessarily affecting day 90 mRS. As recommended, every component of the composite will be evaluated as secondary outcome measure.(17) Medical events will be monitored until day 7 because perioperative complications used to happen during the first week as reported by different trials. (18-20)

- For the secondary outcomes they include successful recanalization but there is no mention of infarct volume – an important variable in terms of outcome.

Successful recanalization will be used as a secondary classical thrombectomy outcome measure. We agree that infarct volume could be an important surrogate endpoint but it will not be evaluated in our study. Notably, radiological standardization of this parameter is difficult in a multicenter trial.

- There is no definition of what constitutes hypotension – how low and how long?

Concerning arterial pressure, investigator will have to treat every episode of systolic blood pressure < 140mmHg with fluids and/or vasopressors. Hypotension will be monitored as a secondary outcome

measure. Indeed, we apologize because its definition is not specified. As defined by others, one episode of systolic blood pressure < 120 mmHg during the prespecified time points of blood pressure measurement will be considered hypotension.(2) We added this definition in Supplementary file 1: AMETIS trial data collection.

- There is not enough information on the specific important time points which will affect outcome. Door to groin puncture time – is this hospital door or radiology suite? Stroke onset time? Time of induction of anaesthesia / sedation, duration of procedure?

We apologize for the lack of information concerning important time points. We pointed out time points definitions and clearly notified the different evaluable time delays:

- o Stroke onset to door delay is time from stroke symptom (or last time seen well for wake-up strokes) to actual hospital admission
- o Door to groin puncture delay is time from actual hospital admission to groin puncture
- o Stroke onset to groin puncture delay is time from stroke symptom (or last time seen well for wake-up strokes) to groin puncture
- o Door to reperfusion delay is time from actual hospital admission to reperfusion
- o GA/CS induction to groin puncture delay is time from administration of the first anaesthetic/sedative agent to groin puncture
- o Duration of the procedure is time from groin puncture to end of procedure (defined as the last set of radiological images)
- o Stroke onset to reperfusion delay is time from stroke symptom (or last time seen well for wake-up strokes) to reperfusion (if any)

We added these elements in secondary outcomes measures of the protocol with definitions in supplementary file 1: AMETIS trial data collection. We also updated supplementary file 2: AMETIS trial statistical analysis plan.

- Why day 7 for medical complications and unexpected ICU admission

Unexpected ICU admission by day 7 will be evaluated as a secondary outcome associated with perioperative complications. Unscheduled rather than scheduled ICU admission reflects complication associated with life-threatening organ dysfunction. As discussed previously, severe complications usually occur within the first week following intervention/anaesthesia. Admission to ICU after this timepoint might capture complications not directly related to the intervention period.

- For sample size estimation, the authors have used 5 previous studies on anterior circulation stroke. However, all of these used a primary outcome of mRS and so they have extrapolated composite scores. I am not sure how reliable this will be.

We used the results of 5 recent trials about thrombectomy in anterior circulation AIS for sample size estimation of the composite outcome. Effectively, these trials used mRS as a primary outcome. As mentioned in prior justification, we choose a composite outcome in order to reflect best the effect of the anaesthetic intervention on significant medical events. Periprocedural complications were reported as secondary outcome measures in these trials and we were then able to optimize sample size estimation.

REVIEWER # 2: We would like to thank Dr Ilyas for the relevant criticisms regarding our manuscript.

- The authors have outlined a protocol designed to determine outcome of endovascular mechanical thrombectomy for anterior circulation acute ischemic stroke under two different anesthesia modalities.

The composite primary outcome included medical complications, though I believe these are both medical and surgical in nature.

We agree that the composite primary outcome integrates in fact medical and surgical complications. We then specified using the term « perioperative complications" that seems to fit best. Modifications were done in the manuscript and in supplementary files.

• Among these complications, pneumonia is listed. Patient may have pre-morbid conditions that predispose them to hospital acquired pneumonia (e.g. chronic obstructive pulmonary disease), or they may present to the hospital with pneumonia. These need to be accounted, and are not listed within the data collection parameters (Supplementary File 1: At randomization).

Concerning pneumonia, these frequently frail patients could indeed have predispositions at hospital admission. As advised, we added “Chronic Obstructive Pulmonary Disease” and “pneumonia at admission” in the data collection parameters in Supplementary file 1.

• Furthermore, the definition of cardiogenic acute pulmonary edema and how it is distinguished from pulmonary edema secondary to fluid overload is not stated clearly as patients with poor ejection fracture are likely to develop pulmonary edema secondary to fluid overload. This outcome is too nuanced and should be removed or clearly defined.

Concerning cardiogenic acute pulmonary edema, as referenced, we used European Society of Anaesthesia / European Society of Intensive Care Medicine definition about outcome measure for clinical effectiveness research in perioperative medicine that is: “evidence of fluid accumulation in the alveoli due to poor cardiac function.” In these recommendations, even if it could be pertinent, no distinction was made with fluid overload.(21) As stated in “supplementary file 1, intraoperative anaesthetic data”, we planned to monitor fluid volume.

• In the Abstract, Introduction: EMT is the standard of care for anterior circulation AIS secondary to emergent large vessel occlusion in patients who qualify (i.e. not all types of anterior circulation AIS). Overall, the syntax and word choice can be improved, for example:

1. In the Abstract: Introduction, the second sentence ("To ensure patient ... proposed.") is awkward.

As requested, we did the modifications in abstract introduction: “Endovascular thrombectomy is the standard of care for anterior circulation acute ischemic stroke (AIS) secondary to emergent large vessel occlusion in patients who qualify. General Anaesthesia (GA) or Conscious Sedation (CS) are usually required to ensure patient comfort and avoid agitation and movement during thrombectomy. However, the question of whether the use of GA or CS might influence functional outcomes remains debated. Indeed, conflicting results exist between observational studies with better outcomes associated with CS and small monocentric randomized controlled trials favouring GA. Therefore, we aim to evaluate the effect of CS versus GA on functional outcome and peri-procedural complications in endovascular mechanical thrombectomy for anterior circulation AIS.”

• 2. In the Introduction: Background and rationale, the sentence ("Notably, anesthetic management ... general anesthesia (GA).") is poorly written and does not effectively convey equipoise between to anesthetic modalities.

As advised, we modified the sentence in Introduction: “Notably, the optimal management strategy during thrombectomy, using either General Anaesthesia (GA) or Conscious Sedation (CS), remains controversial.”

REVIEWER # 3:

• The two primary outcomes are described several times but I am still not quite sure how they are defined and how they will be treated. It seems that "Functional independence" is an ordinal variable which can take the values 0,1 and 2, and the analysis of this is presumably described on page 17, line 31 ff. A Chi2 test or Fishers exact test seems appropriate, but where does the Poisson distribution come from? There is no count variable here...? Or does this refer to the second primary outcome, called "absence of medical complications..."? Perhaps the medical complications are being counted? Although from page 6, it seems that this variable is thought to be dichotomous. On page 34, the Poisson analysis is mentioned again, this time as an analysis of "rate data", what is that? Even so, results will be expressed as relative risks, a term that relate to logistic regression. I am confused.

We thank the reviewer for the relevant comment and we apologize if Statistical Section was not sufficiently clear. It is right that the primary outcome is a dichotomous composite parameter. So, the univariate analysis will be performed using Chi2 or Fishers exact tests. In multivariable analyses,

binary outcomes are commonly analysed by applying a logistic regression model to obtain odds ratios for comparing groups with different sets of characteristics. Although this is often appropriate, there may be situations in which it is more desirable to estimate a relative risk (RR) instead of an odds ratio (OR). Several articles in medical and public health literature point out that when the outcome event is common (incidence of 10% or more), it is often more desirable to estimate an RR since there is an increasing differential between the RR and OR with increasing incidence rates, and there is a tendency for some to interpret ORs as if they are RRs,(22) even if there are some who hold the opinion that the OR should be used even when the outcome is common.(23) RR is easy to interpret and explain, and can be estimated from OR: RR = OR/[(1-probability in reference group) + (probability in reference group x OR)]. (24) However, adjustment for confounding (in multivariable analyses) is still on the OR scale and confidence intervals are too narrow. Also, RR regression is preferred as it allows the direct estimation of RR. There are several options for how to estimate RRs, for example using log-binomial regression model. Zou describes a method to calculate relative risks using Poisson regression with a robust error variance; which corresponds to the method that we propose to perform.(25)

- The sample size calculations are OK, although a 10% loss to follow-up will leave only 270*0.9=243 patients, i.e. not quite 124 in each group. Also, the sample size may not be sufficient for the adjusted analyses.

We thank the reviewer for the helpful comment. We agree. We propose to modify “Assuming lost to follow-up and modified intention to treat population requirements (as defined in supplementary file 2) between 5% and 10%”.

- The covariates are described adequately.

Thanks for the comment.

- In general, I prefer the term "multiple regression" to "multivariable analysis" or just mentioning that the models contain multiple covariates.

We agree. The revised manuscript was modified accordingly.

1. Peng Y, Li Y, Jian M, et al. Choice of ANesthesia for EndoVAScular Treatment of Acute Ischemic Stroke: Protocol for a randomized controlled (CANVAS) trial. International journal of stroke : official journal of the International Stroke Society 2017; 12(9): 991-7.
2. Schonenberger S, Uhlmann L, Hacke W, et al. Effect of Conscious Sedation vs General Anesthesia on Early Neurological Improvement Among Patients With Ischemic Stroke Undergoing Endovascular Thrombectomy: A Randomized Clinical Trial. Jama 2016; 316(19): 1986-96.
3. Goyal M, Demchuk AM, Menon BK, et al. Randomized assessment of rapid endovascular treatment of ischemic stroke. The New England journal of medicine 2015; 372(11): 1019-30.
4. Berkhemer OA, Fransen PS, Beumer D, et al. A randomized trial of intraarterial treatment for acute ischemic stroke. The New England journal of medicine 2015; 372(1): 11-20.
5. Campbell BC, Mitchell PJ, Kleinig TJ, et al. Endovascular therapy for ischemic stroke with perfusion-imaging selection. The New England journal of medicine 2015; 372(11): 1009-18.
6. Saver JL, Goyal M, Bonafe A, et al. Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke. The New England journal of medicine 2015; 372(24): 2285-95.
7. Jovin TG, Chamorro A, Cobo E, et al. Thrombectomy within 8 hours after symptom onset in ischemic stroke. The New England journal of medicine 2015; 372(24): 2296-306.
8. Bracard S, Ducrocq X, Mas JL, et al. Mechanical thrombectomy after intravenous alteplase versus alteplase alone after stroke (THRACE): a randomised controlled trial. The Lancet Neurology 2016; 15(11): 1138-47.
9. Lowhagen Henden P, Rentzos A, Karlsson JE, et al. General Anesthesia Versus Conscious Sedation for Endovascular Treatment of Acute Ischemic Stroke: The AnStroke Trial (Anesthesia During Stroke). Stroke 2017; 48(6): 1601-7.
10. Powers WJ, Rabinstein AA, Ackerson T, et al. 2018 Guidelines for the Early Management of Patients With Acute Ischemic Stroke: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. Stroke 2018.

11. Nogueira RG, Jadhav AP, Haussen DC, et al. Thrombectomy 6 to 24 Hours after Stroke with a Mismatch between Deficit and Infarct. The New England journal of medicine 2018; 378(1): 11-21.
12. Rivera-Lara L, Zorrilla-Vaca A, Geocadin RG, Healy RJ, Ziai W, Mirski MA. Cerebral Autoregulation-oriented Therapy at the Bedside: A Comprehensive Review. Anesthesiology 2017; 126(6): 1187-99.
13. Nordstrom CH, Koskinen LO, Olivecrona M. Aspects on the Physiological and Biochemical Foundations of Neurocritical Care. Frontiers in neurology 2017; 8: 274.
14. Abbott TEF, Pearse RM. Depth of Anesthesia and Postoperative Delirium. Jama 2019; 321(5): 459-60.
15. Simonsen CZ, Yoo AJ, Sorensen LH, et al. Effect of General Anesthesia and Conscious Sedation During Endovascular Therapy on Infarct Growth and Clinical Outcomes in Acute Ischemic Stroke: A Randomized Clinical Trial. JAMA neurology 2018.
16. Wildes TS, Mickle AM, Ben Abdallah A, et al. Effect of Electroencephalography-Guided Anesthetic Administration on Postoperative Delirium Among Older Adults Undergoing Major Surgery: The ENGAGES Randomized Clinical Trial. Jama 2019; 321(5): 473-83.
17. Irony TZ. The "Utility" in Composite Outcome Measures: Measuring What Is Important to Patients. Jama 2017; 318(18): 1820-1.
18. Futier E, Constantin JM, Paugam-Burtz C, et al. A trial of intraoperative low-tidal-volume ventilation in abdominal surgery. The New England journal of medicine 2013; 369(5): 428-37.
19. Fernandez-Bustamante A, Frendl G, Sprung J, et al. Postoperative Pulmonary Complications, Early Mortality, and Hospital Stay Following Noncardiothoracic Surgery: A Multicenter Study by the Perioperative Research Network Investigators. JAMA Surg 2017; 152(2): 157-66.
20. Group PS, Devereaux PJ, Yang H, et al. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. Lancet 2008; 371(9627): 1839-47.
21. Jammer I, Wickboldt N, Sander M, et al. Standards for definitions and use of outcome measures for clinical effectiveness research in perioperative medicine: European Perioperative Clinical Outcome (EPCO) definitions: a statement from the ESA-ESICM joint taskforce on perioperative outcome measures. European journal of anaesthesiology 2015; 32(2): 88-105.
22. Greenland S. Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. American journal of epidemiology 2004; 160(4): 301-5.
23. Cook TD. Advanced statistics: up with odds ratios! A case for odds ratios when outcomes are common. Acad Emerg Med 2002; 9(12): 1430-4.
24. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. Jama 1998; 280(19): 1690-1.
25. Zou G. A modified poisson regression approach to prospective studies with binary data. American journal of epidemiology 2004; 159(7): 702-6.

VERSION 2 – REVIEW

REVIEWER	Judith Dinsmore St Georges University Hospital, London, UK
REVIEW RETURNED	13-Mar-2019

GENERAL COMMENTS	Thank you to the authors for responding to my queries. They have answered the majority of my concerns. I have no further comments.

REVIEWER	Lene Theil Skovgaard Department of Public Health, Section of Biostatistics,

	University of Copenhagen, Denmark
REVIEW RETURNED	15-Mar-2019

GENERAL COMMENTS	It is still not clear from the text that the primary outcome is binary, I only now know that from the author's response. Also, the use of Poisson regression with robust standard errors for this kind of data should be stated clearly. I am a bit worried that this method may not be adequate. In the reference paper by Zou, the method is only tried out for tables, i.e. for categorical covariates, and my concern is that it may not work for quantitative covariates, in the same way that binomial regression with a log-link may fail to converge. In the Binomial regression, the convergence problems come from predicting probabilities above 1, and this is no problem, whereas this is not seen as a problem when using a Poisson distribution. You may still have predictions above 1 (which are of course nonsense), but you will not detect it. Another concern is the number of covariates for the regression models. You should take care not to include too many.

VERSION 2 – AUTHOR RESPONSE

REVIEWER # 1:

Reviewer Name: Judith Dinsmore

Institution and Country: St Georges University Hospital, London, UK

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below

Thank you to the authors for responding to my queries. They have answered the majority of my concerns. I have no further comments.

We thank Doctor Dinsmore.

REVIEWER # 3:

Reviewer Name: Lene Theil Skovgaard

Institution and Country: Department of Public Health,

Section of Biostatistics,

University of Copenhagen,

Denmark

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below

It is still not clear from the text that the primary outcome is binary, I only now know that from the author's response.

Thank you for spotting this imprecision. We have now provided more explicit information on the primary outcome in the revised version of the manuscript.

Also, the use of Poisson regression with robust standard errors for this kind of data should be stated clearly.

We concur that the use of Poisson regression with robust standard errors should have be more clearly stated. The revised version of the manuscript was modified accordingly.

I am a bit worried that this method may not be adequate. In the reference paper by Zou, the method is only tried out for tables, i.e. for categorical covariates, and my concern is that it may not work for quantitative covariates, in the same way that binomial regression with a log-link may fail to converge. In the Binomial regression, the convergence problems come from predicting probabilities above 1, and this is no problem, whereas this is not seen as a problem when using a Poisson distribution. You may still have predictions above 1 (which are of course nonsense), but you will not detect it.

In the paper published by Knol et al. (Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression. CMAJ. 2012), the authors “illustrate the difference between risk ratios and odds ratios using clinical examples, and describe the magnitude of the problem in the literature.” Interestingly, they reviewed available methods to obtain adjusted risk ratios and evaluated these methods by means of simulations, and concluded that “The Mantel–Haenszel risk ratio method, log–binomial regression, Poisson regression with robust standard errors, and the doubling-of-cases method with robust standard errors gave correct risk ratios and confidence intervals.” However, we agree with the Reviewer that particular attention must be paid to the covariates used in multivariable regressions, especially quantitative covariates for which convergence issues can be raised. As presented in the statistical analysis plan, that should only concern “time delays”. In addition, sensitivity analyses considering these covariates will be proposed, dichotomizing according to the statistical distribution and to the clinical relevance. In previous papers by our study group (for example, Futier et al. Effect of Individualized vs Standard Blood Pressure Management Strategies on Postoperative Organ Dysfunction Among High-Risk Patients Undergoing Major Surgery: A Randomized Clinical Trial. JAMA. 2017; Futier et al. A trial of intraoperative low-tidal-volume ventilation in abdominal surgery. N Engl J Med. 2013), adjusted analyses were performed with the use of robust Poisson generalized linear model regression, including a random effect to account for center effect, and are presented as relative risks with 95% confidence intervals. Finally, this statistical approach has also recently been used in two research protocols published in the BMJ Open journal by our group, without leading to methodological concern (Molliex et al. Stepped wedge cluster randomised controlled trial to assess the effectiveness of an optimisation strategy for general anaesthesia on postoperative morbidity and mortality in elderly patients (the OPTI-AGED study): a study protocol. BMJ Open. 2018; Bouvier et al. Assessment of the advantage of the serum S100B protein biomonitoring in the management of pediatric mild traumatic brain injury – PROS100B: protocol of a multicenter unblinded stepped-wedge cluster randomized trial. BMJ Open (in press).

Another concern is the number of covariates for the regression models. You should take care not to include too many.

Thanks for the helpful and relevant comment. Special attention will be given not to include too many covariates.

VERSION 3 – REVIEW

REVIEWER	Lene Theil Skovgaard Department of Public Health, Section of Biostatistics, University of Copenhagen, Denmark
REVIEW RETURNED	20-May-2019

GENERAL COMMENTS	I appreciate that the authors have now included the term "binary" on p. 12, in the section on "Primary outcome measure", but not in the "Primary objective" section on p. 8, even if this contains more or less the same frasing. Neither is it stated on p. 16 in connection with the statistical analysis. Also, very little change has been made regarding the use of Poisson regression for this binary outcome variable. The text includes no justification for not doing a logistic regression, and I guess I will not be the only reader wondering why you turn from a "table analysis" (note here, that Fishers -no c in Fisher- test is always appropriate, so you need not consider chisqure at all) to a Poisson regression. I guess you will want to write something about being able to express the results as risk ratios instead of odds ratios, and that this may be troublesome to do in a logistic regression, because a log-link in the Binomial distribution may give rise to convergence problems. And you should quote the relevant references here.